What can we count in language, and what counts in language acquisition,

cognition, and use?

Nick C. Ellis

University of Michigan

ncellis@umich.edu

"Everything that can be counted does not necessarily count; everything that counts

cannot necessarily be counted." (Albert Einstein)

1   Frequency and Cognition

"Perception is of definite and probable things" (James 1890).

From its very beginnings, psychological research has recognized three major

experiential factors that affect cognition: frequency, recency, and context (e.g., Anderson

2000; Ebbinghaus 1885; Bartlett [1932] 1967). Learning, memory and perception are all

affected by frequency of usage: The more times we experience something, the stronger

our memory for it, and the more fluently it is accessed. The more recently we have

experienced something, the stronger our memory for it, and the more fluently it is

accessed. (Hence your more fluent reading of the prior sentence than the one before). The

more times we experience conjunctions of features, the more they become associated in

our minds and the more these subsequently affect perception and categorization; so a

stimulus becomes associated to a context and we become more likely to perceive it in that

context. The power law of learning (Anderson 1982; Ellis and Schmidt 1998; Newell

1990) describes the relationships between practice and performance in the acquisition of

a wide range of cognitive skills – the greater the practice, the greater the performance,

although effects of practice are largest at early stages of leaning, thereafter diminishing

and eventually reaching asymptote. The power function relating probability of recall and

recency is known as the forgetting curve (Baddeley 1997; Ebbinghaus 1885).

William James' words which begin this section concern the effects of frequency

upon perception. There is a lot more to perception than meets the eye, or ear. A percept is

a complex state of consciousness in which antecedent sensation is supplemented by

consequent ideas which are closely combined to it by association. The cerebral conditions

of the perception of things are thus the paths of association irradiating from them. If a certain sensation is strongly associated with the attributes of a certain thing, that thing is almost sure to be perceived when we get that sensation. But where the sensation is associated with more than one reality, unconscious processes weigh the odds, and we perceive the most probable thing: "all brain-processes are such as give rise to what we may call FIGURED consciousness" (James, 1890, p. 82). Accurate and fluent perception thus rests on the perceiver having acquired the appropriately weighted range of associations for each element of the sensory input.

It is human categorization ability which provides the most persuasive testament to our incessant unconscious figuring or 'tallying' (Ellis 2002). We know that natural categories are fuzzy rather than monothetic. Wittgenstein's (1953) consideration of the concept game showed that no set of features that we can list covers all the things that we call games, ranging as the exemplars variously do from soccer, through chess, bridge, and poker, to solitaire. Instead, what organizes these exemplars into the game category is a set of family resemblances among these members -- son may be like mother, and mother like sister, but in a very different way. And we learn about these families, like our own, from experience. Exemplars are similar if they have many features in common and few distinctive attributes (features belonging to one but not the other); the more similar are two objects on these quantitative grounds, the faster are people at judging them to be similar (Tversky 1977). Prototypes, exemplars which are most typical of a category, are those which are similar to many members of that category and not similar to members of other categories. Again, the operationalisation of this criterion predicts the speed of human categorization performance -- people more quickly classify as birds sparrows (or

other average sized, average colored, average beaked, average featured specimens) than they do birds with less common features or feature combinations like kiwis or penguins (Rosch and Mervis 1975; Rosch et al. 1976). Prototypes are judged faster and more accurately, even if they themselves have never been seen before -- someone who has never seen a sparrow, yet who has experienced the rest of the run of the avian mill, will still be fast and accurate in judging it to be a bird (Posner and Keele 1970). Such effects make it very clear that although people don't go around consciously counting features, they nevertheless have very accurate knowledge of the underlying frequency distributions and their central tendencies. Cognitive theories of categorization and generalization show how schematic constructions are abstracted over less schematic ones that are inferred inductively by the learner in acquisition (Lakoff 1987; Taylor 1998; Harnad 1987). So Psychology is committed to studying these implicit processes of cognition.

2   Frequency and Language Cognition

The last 50 years of Psycholinguistic research has demonstrated language processing to be exquisitely sensitive to usage frequency at all levels of language representation: phonology and phonotactics, reading, spelling, lexis, morphosyntax, formulaic language, language comprehension, grammaticality, sentence production, and syntax (Ellis 2002). Language knowledge involves statistical knowledge, so humans learn more easily and process more fluently high frequency forms and 'regular' patterns which are exemplified by many types and which have few competitors. Psycholinguistic perspectives thus hold that language learning is the implicit associative learning of representations that reflect the probabilities of occurrence of form-function mappings. Frequency is a key determinant of acquisition because 'rules' of language, at all levels of

analysis from phonology, through syntax, to discourse, are structural regularities which emerge from learners' lifetime unconscious analysis of the distributional characteristics of the language input. In James' terms, learners have to FIGURE language out.

It is these ideas which underpin the last 30 years of investigations of language cognition using connectionist and statistical models Christiansen & Chater, 2001; Elman, et al., 1996; Rumelhart & McClelland, 1986), the competition model of language learning and processing (Bates and MacWhinney 1987; MacWhinney 1987, 1997), the investigation of how frequency and repetition bring about form in language and how probabilistic knowledge drives language comprehension and production (Jurafsky and Martin 2000; Ellis 2002; Bybee and Hopper 2001; Jurafsky 2002; Bod, Hay, and Jannedy 2003; Ellis 2002; Hoey 2005), and the proper empirical investigations of the structure of language by means of corpus analysis exemplified in this volume. Corpus linguistics allows us to count the relevant frequencies in the input.

Frequency, learning, and language come together in Usage-based approaches which hold that we learn linguistic constructions while engaging in communication, the "interpersonal communicative and cognitive processes that everywhere and always shape language" (Slobin 1997). Constructions are form-meaning mappings, conventionalized in the speech community, and entrenched as language knowledge in the learner's mind. They are the symbolic units of language relating the defining properties of their morphological, syntactic, and lexical form with particular semantic, pragmatic, and discourse functions (Croft and Cruise 2004; Robinson and Ellis 2008; Goldberg 2003, 2006; Croft 2001; Tomasello 2003; Bates and MacWhinney 1987; Goldberg 1995; Langacker 1987; Lakoff 1987; Bybee 2008). Goldberg's (2006) Construction Grammar

argues that all grammatical phenomena can be understood as learned pairings of form (from morphemes, words, idioms, to partially lexically filled and fully general phrasal patterns) and their associated semantic or discourse functions: "the network of constructions captures our grammatical knowledge *in toto*, i.e. It's constructions all the way down" (Goldberg 2006, p. 18). Such beliefs, increasingly influential in the study of child language acquisition, have turned upside down generative assumptions of innate language acquisition devices, the continuity hypothesis, and top-down, rule-governed, processing, bringing back data-driven, emergent accounts of linguistic systematicities. Constructionist theories of child language acquisition use dense longitudinal corpora to chart the emergence of creative linguistic competence from children's analyses of the utterances in their usage history and from their abstraction of regularities within them (Tomasello 1998, 2003; Goldberg 2006, 1995, 2003). Children typically begin with phrases whose verbs are only conservatively extended to other structures. A common developmental sequence is from formula to low-scope slot-and-frame pattern, to creative construction.

3   Frequency and Second Language Cognition

What of second language acquisition (L2A)? Language learners, L1 and L2 both, share the goal of understanding language and how it works. Since they achieve this based upon their experience of language usage, there are many commonalities between first and second language acquisition that can be understood from corpus analyses of input and cognitive- and psycho- linguistic analyses of construction acquisition following associative and cognitive principles of learning and categorization. Therefore Usage-based approaches, Cognitive Linguistics, and Corpus Linguistics are increasingly

influential in L2A research too (Ellis 1998, 2003; Ellis and Cadierno 2009; Collins and Ellis 2009; Robinson and Ellis 2008), albeit with the twist that since they have previously devoted considerable resources to the estimation of the characteristics of another language -- the native tongue in which they have considerable fluency – L2 learners' computations and inductions are often affected by transfer, with L1-tuned expectations and selective attention (Ellis 2006) blinding the acquisition system to aspects of the L2 sample, thus biasing their estimation from naturalistic usage and producing the limited attainment that is typical of adult L2A. Thus L2A is different from L1A in that it involves processes of construction and *re*construction

4   Construction Learning as Associative Learning from Usage

If constructions as form-function mappings are the units of language, then language acquisition involves inducing these associations from experience of language usage. Constructionist accounts of language acquisition thus involve the distributional analysis of the language stream and the parallel analysis of contingent perceptual activity, with abstract constructions being learned from the conspiracy of concrete exemplars of usage following statistical learning mechanisms (Christiansen and Chater 2001) relating input and learner cognition. Psychological analyses of the learning of constructions as form-meaning pairs is informed by the literature on the associative learning of cue-outcome contingencies where the usual determinants include: factors relating to the form such as frequency and salience; factors relating to the interpretation such as significance in the comprehension of the overall utterance, prototypicality, generality, and redundancy; factors relating to the contingency of form and function; and factors relating to learner attention, such as automaticity, transfer, overshadowing, and blocking (Ellis

2002, 2003, 2006, 2008). These various psycholinguistic factors conspire in the acquisition and use of any linguistic construction.

These determinants of learning can be usefully categorized into factors relating to (1) input frequency (type-token frequency, Zipfian distribution, recency), (2) form (salience and perception), (3) function (prototypicality of meaning, importance of form for message comprehension, redundancy), and (4) interactions between these (contingency of form-function mapping). The following subsections briefly consider each in turn, along with studies demonstrating their applicability:

4.1 Input frequency (construction frequency, type-token frequency, Zipfian distribution, recency)

4.1.1 Construction frequency

Frequency of exposure promotes learning. Ellis' (2002a) review illustrates how frequency effects the processing of phonology and phonotactics, reading, spelling, lexis, morphosyntax, formulaic language, language comprehension, grammaticality, sentence production, and syntax. That language users are sensitive to the input frequencies of these patterns entails that they must have registered their occurrence in processing. These frequency effects are thus compelling evidence for usage-based models of language acquisition which emphasize the role of input.

4.1.2 Type and token frequency

Token frequency counts how often a particular form appears in the input. Type frequency, on the other hand, refers to the number of distinct lexical items that can be substituted in a given slot in a construction, whether it is a word-level construction for

inflection or a syntactic construction specifying the relation among words. For example, the "regular" English past tense -ed has a very high type frequency because it applies to thousands of different types of verbs, whereas the vowel change exemplified in swam and rang has much lower type frequency. The productivity of phonological, morphological, and syntactic patterns is a function of type rather than token frequency (Bybee and Hopper 2001). This is because: (a) the more lexical items that are heard in a certain position in a construction, the less likely it is that the construction is associated with a particular lexical item and the more likely it is that a general category is formed over the items that occur in that position; (b) the more items the category must cover, the more general are its criterial features and the more likely it is to extend to new items; and (c) high type frequency ensures that a construction is used frequently, thus strengthening its representational schema and making it more accessible for further use with new items (Bybee and Thompson 2000). In contrast, high token frequency promotes the entrenchment or conservation of irregular forms and idioms; the irregular forms only survive because they are high frequency. These findings support language's place at the center of cognitive research into human categorization, which also emphasizes the importance of type frequency in classification.

Such effects are extremely robust in the dynamics of language usage and structural evolution: (1) For token frequency, entrenchment, and protection from change, Pagel, Atkinson & Meade (2007) used a database of 200 fundamental vocabulary meanings in 87 Indo-European languages to calculate how quickly the different meanings evolved over time. Records of everyday speech in English, Spanish, Russian and Greek showed that high token-frequency words that were used more often in everyday language

evolved more slowly. Across all 200 meanings, word token frequency of usage determined their rate of replacement over thousands of years, with the most commonly-used words, such as numbers, changing very little. (2) For type and token frequency, and the effects of friends and enemies in the dynamics of productivity of patterns in language evolution, Lieberman, Michel, Jackson, Tang, and Nowak (2007) studied the regularization of English verbs over the past 1,200 years. English's proto-Germanic ancestor used an elaborate system of productive conjugations to signify past tense whereas Modern English makes much more productive use of the dental suffix, '-ed'. Lieberman at al. chart the emergence of this linguistic rule amidst the evolutionary decay of its exceptions. By tracking inflectional changes to 177 Old-English irregular verbs of which 145 remained irregular in Middle English and 98 are still irregular today, they showed how the rate of regularization depends on the frequency of word usage. The half-life of an irregular verb scales as the square root of its usage frequency: a verb that is 100 times less frequent regularizes 10 times as fast.

4.1.3 Zipfian distribution

Zipf's law states that in human language, the frequency of words decreases as a power function of their rank in the frequency table. If p_f is the proportion of words whose frequency in a given language sample is f, then $p_f \sim f^{-\beta}$, with $\beta \approx 1$. Zipf (1949) showed this scaling relation holds across a wide variety of language samples. Subsequent research has shown that many language events (e.g., frequencies of phoneme and letter strings, of words, of grammatical constructs, of formulaic phrases, etc.) across scales of analysis follow this law (Ferrer i Cancho and Solé 2001, 2003). It has strong empirical

support as a linguistic universal, and, as I shall argue in the closing section of this chapter, its implications are profound for language structure, use, and acquisition. For present purposes, this section focuses upon acquisition.

In the early stages of learning categories from exemplars, acquisition is optimized by the introduction of an initial, low-variance sample centered upon prototypical exemplars (Elio and Anderson 1981, 1984). This low variance sample allows learners to get a fix on what will account for most of the category members. The bounds of the category are defined later by experience of the full breadth of exemplar types. Goldberg, Casenhiser & Sethuraman (2004) demonstrated that in samples of child language acquisition, for a variety of verb-argument constructions (VACs), there is a strong tendency for one single verb to occur with very high frequency in comparison to other verbs used, a profile which closely mirrors that of the mothers' speech to these children. In natural language, Zipf's law (Zipf 1935) describes how the highest frequency words account for the most linguistic tokens. Goldberg et al. (2004) show that Zipf's law applies within VACs too, and they argue that this promotes acquisition: tokens of one particular verb account for the lion's share of instances of each particular argument frame; this pathbreaking verb also is the one with the prototypical meaning from which the construction is derived (see also Ninio 1999, 2006).

Ellis and Ferreira-Junior (2009, 2009) investigate effects upon naturalistic second language acquisition of type/token distributions in the islands comprising the linguistic form of English verb-argument constructions (VACs: VL verb locative, VOL verb object locative, VOO ditransitive) in the ESF corpus (Perdue, 1993). They show that in the naturalistic L2A of English, VAC verb type/token distribution in the input is Zipfian and

learners first acquire the most frequent, prototypical and generic exemplar (e.g. *put* in VOL, *give* in VOO, etc.). Their work further illustrates how acquisition is affected by the frequency and frequency distribution of exemplars within each island of the construction (e.g. [Subj V Obj Obl$_{path/loc}$]), by their prototypicality, and, using a variety of psychological (Shanks 1995) and corpus linguistic association metrics (Gries and Stefanowitsch 2004; Stefanowitsch and Gries 2003), by their contingency of form-function mapping. Ellis and Larsen-Freeman (2009) describe computational (Emergent connectionist) serial-recurrent network models of these various factors as they play out in the emergence of constructions as generalized linguistic schema from their frequency distributions in the input.

This fundamental claim that Zipfian distributional properties of language usage helps to make language learnable has thus begun to be explored for these three verb argument constructions, at least. It remains an important corpus linguistic research agenda to explore its generality across the wide range of the constructicon.

4.1.4   Recency

Language processing also reflects recency effects. This phenomenon, known as priming, may be observed in phonology, conceptual representations, lexical choice, and syntax (Pickering and Ferreira 2008). Syntactic priming refers to the phenomenon of using a particular syntactic structure given prior exposure to the same structure. This behavior has been observed when speakers hear, speak, read or write sentences (Bock 1986; Pickering 2006; Pickering and Garrod 2006). For L2A, Gries and Wulff (2005) showed (i) that advanced L2 learners of English showed syntactic priming for ditransitive (e.g., *The racing driver showed the helpful mechanic*) and prepositional dative (e.g., *The*

racing driver showed the torn overall …) argument structure constructions in a sentence completion task, (ii) that their semantic knowledge of argument structure constructions affected their grouping of sentences in a sorting task, and (iii) that their priming effects closely resembled those of native speakers of English in that they were very highly correlated with native speakers' verbal subcategorization preferences whilst completely uncorrelated with the subcategorization preferences of the German translation equivalents of these verbs. There is now a growing body of research demonstrating such L2 syntactic priming effects (McDonough 2006; McDonough and Mackey 2006; McDonough and Trofimovich 2008)

4.2 Form (salience and perception)

The general perceived strength of stimuli is commonly referred to as their salience. Low salience cues tend to be less readily learned. Ellis (2006, 2006) summarized the associative learning research demonstrating that selective attention, salience, expectation, and surprise are key elements in the analysis of all learning, animal and human alike. As the Rescorla-Wagner (1972) model encapsulates, the amount of learning induced from an experience of a cue-outcome association depends crucially upon the salience of the cue and the importance of the outcome.

Many grammatical meaning-form relationships, particularly those that are notoriously difficult for second language learners like grammatical particles and inflections such as the third person singular *–s* of English, are of low salience in the language stream. For example, some forms are more salient: '*today*' is a stronger psychophysical form in the input than is the morpheme '-*s*' marking 3rd person singular present tense, thus while both provide cues to present time, *today* is much more likely to

be perceived, and *-s* can thus become overshadowed and blocked, making it difficult for second language learners of English to acquire (Ellis 2006, 2008; Goldschneider and DeKeyser 2001).

4.3   Function (prototypicality of meaning, importance of form for message comprehension, redundancy)

4.3.1   Prototypicality of meaning

Some members of categories are more typical of the category than others – they show the family resemblance more clearly. In the prototype theory of concepts (Rosch and Mervis 1975; Rosch et al. 1976), the prototype as an idealized central description is the best example of the category, appropriately summarizing the most representative attributes of a category. As the typical instance of a category, it serves as the benchmark against which surrounding, less representative instances are classified. The greater the token frequency of an exemplar, the more it contributes to defining the category, and the greater the likelihood it will be considered the prototype. The best way to teach a concept is to show an example of it. So the best way to introduce a category is to show a prototypical example. Ellis & Ferreira-Junior (2009) show that the verbs that second language learners first used in particular VACs are prototypical and generic in function (*go* for VL, *put* for VOL, and *give* for VOO). The same has been shown for child language acquisition, where a small group of semantically general verbs, often referred to as *light verbs* (e.g., *go, do, make, come*) are learned early (Clark 1978; Ninio 1999; Pinker 1989). Ninio argues that, because most of their semantics consist of some schematic notion of transitivity with the addition of a minimum specific element, they are semantically suitable, salient, and frequent; hence, learners start transitive word

combinations with these generic verbs. Thereafter, as Clark describes, "many uses of these verbs are replaced, as children get older, by more specific terms. . . . General purpose verbs, of course, continue to be used but become proportionately less frequent as children acquire more words for specific categories of actions" (p. 53).

4.3.2   Redundancy

The Rescorla-Wagner model (1972) also summarizes how redundant cues tend not to be acquired. Not only are many grammatical meaning-form relationships low in salience, but they can also be redundant in the understanding of the meaning of an utterance. For example, it is often unnecessary to interpret inflections marking grammatical meanings such as tense because they are usually accompanied by adverbs that indicate the temporal reference. Second language learners' reliance upon adverbial over inflectional cues to tense has been extensively documented in longitudinal studies of naturalistic acquisition (Dietrich, Klein, and Noyau 1995; Bardovi-Harlig 2000), training experiments (Ellis 2007; Ellis and Sagarra 2010), and studies of L2 language processing (Van Patten 2006; Ellis and Sagarra 2010).

4.4   Interactions between these (contingency of form-function mapping)

Psychological research into associative learning has long recognized that while frequency of form is important, so too is contingency of mapping (Shanks 1995). Consider how, in the learning of the category of birds, while eyes and wings are equally frequently experienced features in the exemplars, it is wings which are distinctive in differentiating birds from other animals. Wings are important features to learning the category of birds because they are reliably associated with class membership, eyes are neither. Raw frequency of occurrence is less important than the contingency between cue

and interpretation. Distinctiveness or reliability of form-function mapping is a driving force of all associative learning, to the degree that the field of its study has been known as 'contingency learning' since Rescorla (1968) showed that for classical conditioning, if one removed the contingency between the conditioned stimulus (CS) and the unconditioned (US), preserving the temporal pairing between CS and US but adding additional trials where the US appeared on its own, then animals did not develop a conditioned response to the CS. This result was a milestone in the development of learning theory because it implied that it was contingency, not temporal pairing, that generated conditioned responding. Contingency, and its associated aspects of predictive value, cue validity, information gain, and statistical association, have been at the core of learning theory ever since. It is central in psycholinguistic theories of language acquisition too (Ellis 2008; MacWhinney 1987; Ellis 2006, 2006; Gries and Wulff 2005), with the most developed account for second language acquisition being that of the Competition model (MacWhinney 1987, 1997, 2001). Ellis and Ferreira-Junior (2009) use ΔP and collostructional analysis measures (Gries and Stefanowitsch 2004; Stefanowitsch and Gries 2003) to investigate effects of form-function contingency upon L2 VAC acquisition. Wulff, Ellis, Römer, Bardovi-Harlig and LeBlanc (2009) use multiple distinctive collexeme analysis to investigate effects of reliability of form-function mapping in the second language acquisition of tense and aspect. Boyd and Goldberg (Boyd and Goldberg 2009) use conditional probabilities to investigate contingency effects in VAC acquisition. This is still an active area of inquiry, and more research is required before we know which statistical measures of form-function contingency are more predictive of acquisition and processing.

4.5   The Many Aspects of Frequency and their Research Consequences

This section has gathered a range of frequency-related factors that influence the acquisition of linguistic constructions:

1.   the frequency, the frequency distribution, and the salience of the form types,

2.   the frequency, the frequency distribution, the prototypicality and generality of the semantic types, their importance in interpreting the overall construction,

3.   the reliabilities of the mapping between 1 and 2,

4.   the degree to which the different elements in the construction sequence (such as the Subj V Obj and Obl islands in the archipelago of the VL verb argument construction) are mutually informative and form predictable chunks.

There are many factors involved, and research to date has tended to look at each hypothesis by hypothesis, variable by variable, one at a time. But they interact. And what we really want is a model of usage and its effects upon acquisition. We can measure these factors individually. But such counts are vague indicators of how the demands of human interaction affect the content and ongoing co-adaptation of discourse, how this is perceived and interpreted, how usage episodes are assimilated into the learner's system, and how the system reacts accordingly. We need theoretical models of learning, development, and emergence that takes these factors into account dynamically. I will return to this prospect in sections 7-8 after first considering some implications for instruction.

5   Language Learning as Estimation from Sample: Implications for Instruction

Language learners have limited experience of the target language. Their limited exposure poses them the task of estimating how linguistic constructions work from an input sample that is incomplete, uncertain, and noisy. Native-like fluency, idiomaticity, and selection are another level of difficulty again. For a good fit, every utterance has to be chosen, from a wide range of possible expressions, to be appropriate for that idea, for that speaker, for that place, and for that time. And again, learners can only *estimate* this from their finite experience.

Like other estimation problems, successful determination of the population characteristics is a matter of statistical sampling, description, and inference. There are three fundamental instructional aspects of this conception of language learning as statistical sampling and estimation, and Corpus Linguistics is central in each.

5.1   Sample Size

The first and foremost concerns sample size: As in all surveys, the bigger the sample, the more accurate the estimates, but also the greater the costs. Native speakers estimate their language over a lifespan of usage. L2 and foreign language learners just don't have that much time or resource. Thus, they are faced with a task of optimizing their estimates of language from a limited sample of exposure.

Corpus Linguistic analyses are essential to the determination of which constructions of differing degrees of schematicity are worthy of instruction, their relative frequency, and their best (= prototypical and most frequent) examples for instruction and assessment. Gries (2008) describes how three basic methods of corpus linguistics

(frequency lists, concordances, and collocations) inform the instruction of second language constructions.

5.2   Sample Selection

Principles of survey design dictate that a sample must properly represent the strata of the population of greatest concern. Corpus linguistics, genre analysis, and needs analysis have a large role to play in identifying the linguistic constructions of most relevance to particular learners. For example, every genre of English for Academic Purposes and English for Special Purposes has its own phraseology, and learning to be effective in the genre involves learning this (Swales 1990). Lexicographers develop their learner dictionaries upon relevant corpora and dictionaries focus upon examples of usage as much as definitions, or even more so. Good grammars are now frequency informed. Corpus linguistic analysis techniques have been used to identify the words relevant to academic English (the Academic Word List, Coxhead 2000) and this, together with knowledge of lexical acquisition and cognition, informs vocabulary instruction programs (Nation 2001). Similarly, corpus techniques have been used to identify formulaic phrases that are of special relevance to academic discourse and to inform their instruction (the Academic Formulas List, Ellis, Simpson-Vlach, and Maynard 2008).

5.3   Sample Sequencing

Corpus linguistics also has a role to play in informing the ordering of exemplars for optimal acquisition of a schematic construction. The research reviewed above suggests that an initial, low-variance sample centered upon prototypical exemplars allows learners to get a 'fix' on the central tendency of a schematic construction, and then the

introduction of more diverse exemplars facilitates learners to determine the full range and bounds of the category. Although, as explained in section 4.1.3, there is work to-be-done on determining its applicability to particular constructions, and particular learners and their L1s, in second language acquisition, this is probably a generally useful instructional heuristic. Readings in Robinson and Ellis (2008) show how an understanding of the item-based nature of construction learning inspires the creation and evaluation of instructional tasks, materials, and syllabi, and how cognitive linguistic analyses can be used to inform learners how constructions are conventionalized ways of matching certain expressions to specific situations and to guide instructors in isolating and presenting the various conditions that motivate speaker choice.

6   Exploring what counts

Usage is rich in latent linguistic structure, thus frequencies of usage count in the emergence of linguistic constructions. Corpus Linguistics provides the proper empirical means whereby everything in language texts can be counted. But, following the quotation from Einstein that opened this chapter, not everything that we can count in language counts in language cognition and acquisition. If it did, the English articles *the* and *a* alongside frequent morphological inflections would be among the first learned English constructions, rather than the most problematic in L2A.

The evidence gathered so far in this chapter shows clearly that the study of language from corpus linguistic perspectives is a two-limbed stool without triangulation from an understanding of the psychology of cognition, learning, attention, and development. Sensation is not perception, and the psychophysical relations mapping

physical onto psychological scales are complex. The world of conscious experience is not the world itself but a perception crucially determined by attentional limitations, prior knowledge, and context. Not every experience is equal – effects of practice are greatest at early stages but eventually reach asymptote. The associative learning of constructions as form-meaning pairs is affected by: factors relating to the form such as frequency and salience; factors relating to the interpretation such as significance in the comprehension of the overall utterance, prototypicality, generality, and redundancy; factors relating to the contingency of form and function; and factors relating to learner attention, such as automaticity, transfer, and blocking.

We need models of usage and its effects upon acquisition. Univariate counts are vague indicators of how the demands of human interaction affect the content and ongoing co-adaptation of discourse, how this is perceived and interpreted, how usage episodes are assimilated into the learner's system, and how the linguistic system reacts accordingly. We need models of learning, development, and emergence that take all these factors into account dynamically.

7   Emergentism and Complexity

Although the above conclusion is not contentious, the proper path to its solution is more debatable. In these final sections of my introductory review, I outline Emergentist and related approaches that I believe to be useful in guiding future research. Two key motivations of the editors of this volume are those of empirical rigor and interdisciplinarity. Emergentism fits well, I believe, as a general framework in that it is

as quantitative as anything we have considered here so far, but more so in its recognition of multivariate, multi-agent, often non-linear, interactions.

Language usage involves agents and their processes at many levels, from neuron through self, to society. We need to try to understand language emergence as a function of interactions within and between them. This is a tall order. Hence Saussure's observation that "to speak of a 'linguistic law' in general is like trying to lay hands on a ghost... Synchronic laws are general, but not imperative.. [they] are imposed upon speakers by the constraints of common usage... In short, when one speaks of a synchronic law, one is speaking of an arrangement, or a principle of regularity" (Saussure 1916). Nevertheless, 100 years of subsequent work psycholinguistics has put substantial flesh on the bone. And more recently, work within Emergentism, Complex Adaptive Systems (CAS) , and Dynamic Systems Theory (DST) has started to describe a number of scale-free, domain-general processes which characterize the emergence of pattern across the physical, natural, and social world:

Emergentism and Complexity Theory (MacWhinney 1999; Ellis 1998; Elman et al. 1996; Larsen-Freeman 1997; Larsen-Freeman and Cameron 2008; Ellis and Larsen-Freeman 2009, 2006) analyze how complex patterns emerge from the interactions of many agents, how each emergent level cannot come into being except by involving the levels that lie below it, and how at each higher level there are new and emergent kinds of relatedness not found below: "More is different" (Anderson 1972). These approaches align well with DST which considers how cognitive, social and environmental factors are in continuous interactions, where flux and individual variation abound, and where cause-effect relationships are non-linear, multivariate and interactive in time (Ellis and Larsen-

Freeman 2006, 2006; van Geert 1991; Port and Van Gelder 1995; Spivey 2006; de Bot, Lowie, and Verspoor 2007; Spencer, Thomas, and McClelland 2009; Ellis 2008).

"Emergentists believe that simple learning mechanisms, operating in and across the human systems for perception, motor-action and cognition as they are exposed to language data as part of a communicatively-rich human social environment by an organism eager to exploit the functionality of language, suffice to drive the emergence of complex language representations." (Ellis 1998, p. 657). Language cannot be understood in neurological or physical terms alone, nevertheless, neurobiology and physics play essential roles in the complex interrelations; equally from the top down, though language cannot be understood purely from introspection, nevertheless, conscious experience is an essential part too.

Language considered as a CAS of dynamic usage and its experience involves the following key features:

- The system consists of multiple agents (the speakers in the speech community) interacting with one another.

- The system is adaptive, that is, speakers' behavior is based on their past interactions, and current and past interactions together feed forward into future behavior.

- A speaker's behavior is the consequence of competing factors ranging from perceptual mechanics to social motivations.

- The structures of language emerge from interrelated patterns of experience, social interaction, and cognitive processes.

The advantage of viewing language as a CAS is that it provides a unified account of seemingly unrelated linguistic phenomena (Holland 1998, 1995; Beckner et al. 2009).

These phenomena include: variation at all levels of linguistic organization; the probabilistic nature of linguistic behavior; continuous change within agents and across speech communities; the emergence of grammatical regularities from the interaction of agents in language use; and stage-like transitions due to underlying nonlinear processes. Much of CAS research investigates these interactions through the use of computer simulations (Ellis and Larsen-Freeman 2009). One reason to be excited about a CAS / Corpus Linguistics synergy is that the scale-free phenomena that are characteristic of complex systems were indeed first identified in language corpora.

8   Zipf, Corpora, and Complex Adaptive Systems

Zipf's (1935) analyses of frequency patterns in linguistic corpora, however small they might seem in today's terms, allowed him to identify a scaling law that was universal across language usage. He later attributed this law to the Principle of Least Effort, whereby natural languages realize effective communication by balancing speaker effort (optimized by having fewer words to be learned and accessed in speech production) and ambiguity of speech comprehension (minimized by having many words, one for each different meaning) (Zipf 1949). Many language events across scales of analysis follow his power law: phoneme and letter strings (Kello and Beltz 2009), words (Evert 2005), grammatical constructs (Ninio 2006; O'Donnell and Ellis 2010), formulaic phrases (O'Donnell and Ellis 2009), etc. Scale-free laws also pervade language structures, such as scale-free networks in collocation (Solé et al. 2005), in morphosyntactic productivity (Baayen 2008), in grammatical dependencies (Ferrer i Cancho and Solé 2001, 2003; Ferrer i Cancho, Solé, and Köhler 2004), and in networks of speakers, and language dynamics such as in speech perception and production, in language processing,

in language acquisition, and in language change (Ninio 2006; Ellis 2008). Zipfian covering determines basic categorization, the structure of semantic classes, and the language form-semantic structure interface (Tennenbaum 2005; Manin 2008). Language structure and usage are inseparable, and scale-free laws pervade both. And not just language structure and use.

Power law behavior like this has since been shown to apply to a wide variety of structures, networks, and dynamic processes in physical, biological, technological, social, cognitive, and psychological systems of various kinds (e.g. magnitudes of earthquakes, sizes of meteor craters, populations of cities, citations of scientific papers, number of hits received by web sites, perceptual psychophysics, memory, categorization, etc.) (Newman 2005; Kello et al. 2010). It has become a hallmark of Complex Systems theory where so-called fat-tailed distributions characterize phenomena at the edge of chaos, at a self-organized criticality phase-transition point midway between stable and chaotic domains. The description and analysis of the way in which items (nodes) of different types are arranged into systems (networks) through the connections (edges) formed between them is the focus of the growing field of network science. The ubiquity and diversity of the systems best analyzed as networks, from the connection of proteins in yeast cells to the close association between two actors who have never been co-stars, has given the study of network typologies and dynamics a place alongside the study of other physical laws and properties (Albert and Barabasi 2002; Newman 2003). Properties of networks such as the 'small world' phenomenon (short path between any two nodes even in massive networks), scale-free degree distribution, and the notion of 'preferential attachment' (new nodes added to a network tend to connect to already highly-connected nodes) hold for

networks of language events, structures, and users. Zipfian scale-free laws are universal. They are fundamental too, underlying language processing, learnability, acquisition, usage and change (Ferrer i Cancho and Solé 2001, 2003; Ferrer i Cancho, Solé, and Köhler 2004; Solé et al. 2005). Much remains to be understood, but this is a research area worthy of rich investment, where counting should really count.

Frequency is important to language. Systems depend upon regularity. But not only in the many simple ways. Regular as clockwork proves true in many areas of language representation, change, and processing, as this review has demonstrated. But more is different. In section 7, I argued that the study of language from corpus linguistic perspectives is a two-limbed stool without triangulation from an understanding of the psychology of cognition, learning, attention, and development. Even a three limbed stool does not make much sense without an appreciation of its social use. The cognitive neural networks that compute the associations binding linguistic constructions are embodied, attentionally- and socially- gated, conscious, dialogic, interactive, situated, and cultured (Ellis 2008; Beckner et al. 2009; Ellis and Larsen-Freeman 2009; Bergen and Chang 2003). Language usage, social roles, language learning, and conscious experience are all socially situated, negotiated, scaffolded, and guided. They emerge in the dynamic play of social intercourse. All these factors conspire dynamically in the acquisition and use of any linguistic construction. The future lies in trying to understand the component dynamic interactions at all levels, and the consequent emergence of the complex adaptive system of language itself.

9   References

Albert, R. , and A. Barabasi. 2002. Statistical mechanics of complex networks. *Rev. Mod. Phys* 74:47-97.

Anderson, J. R. 1982. Acquisition of cognitive skill. *Psychological Review* 89 (4):369-406.

———. 2000. *Cognitive psychology and its implications (5th ed.)*. New York: W.H. Freeman.

Anderson, P.W. 1972. More is different. *Science* 177:393-396.

Baayen, R. H. 2008. Corpus linguistics in morphology: morphological productivity. In *Corpus Linguistics. An international handbook*, edited by A. Ludeling and M. Kyto. Berlin: Mouton De Gruyter.

Baddeley, A. D. 1997. *Human memory: Theory and practice.* Revised ed. Hove: Psychology Press.

Bardovi-Harlig, K. 2000. *Tense and aspect in second language acquisition: Form, meaning, and use.* Oxford: Blackwell.

Bartlett, F. C. [1932] 1967. *Remembering: A Study in Experimental and Social Psychology*. Cambridge: Cambridge University Press

Bates, E., and B. MacWhinney. 1987. Competition, variation, and language learning. In *Mechanisms of language acquisition*, edited by B. MacWhinney. Hillsdale, NJ: Lawrence Erlbaum Associates.

Beckner, C., R. Blythe, J. Bybee, M. H. Christiansen, W. Croft, N. C. Ellis, J. Holland, J. Ke, D. Larsen-Freeman, and T. Schoenemann. 2009. Language is a complex adaptive system. Position paper. *Language Learning* 59 Supplement 1:1-26.

Bergen, B.K., and N.C. Chang. 2003. Embodied construction grammar in simulation-based language understanding. In *Construction Grammars: Cognitive grounding and theoretical extensions*, edited by J.-O. Östman and M. Fried. Amsterdam/Philadelphia: John Benjamins,.

Bock, J. K. 1986. Syntactic persistence in language production. *Cognitive Psychology* 18:355-387.

Bod, R., J. Hay, and S. Jannedy, eds. 2003. *Probabilistic linguistics*. Cambridge, MA: MIT Press.

Boyd, J. K., and A. E. Goldberg. 2009. Input effects within a constructionist framework. *Modern Language Journal* 93 (2):418-429.

Bybee, J. 2008. Usage-based grammar and second language acquisition. In *Handbook of cognitive linguistics and second language acquisition*, edited by P. Robinson and N. C. Ellis. London: Routledge.

Bybee, J., and P. Hopper, eds. 2001. *Frequency and the emergence of linguistic structure*. Amsterdam: Benjamins.

Bybee, J., and S. Thompson. 2000. Three frequency effects in syntax. *Berkeley Linguistic Society* 23:65-85.

Christiansen, M. H., and N. Chater, eds. 2001. *Connectionist psycholinguistics*. Westport, CO: Ablex.

Clark, E.V. 1978. Discovering what words can do. In *Papers from the parasession on the lexicon, Chicago Linguistics Society April 14-15, 1978*, edited by D. Farkas, W. M. Jacobsen and K. W. Todrys. Chicago: Chicago Linguistics Society.

Collins, L., and N. C. Ellis. 2009. Input and second language construction learning: frequency, form, and function. *Modern Language Journal* 93 (2):Whole issue.

Coxhead, A. 2000. A new Academic Word List. *TESOL Quarterly* 34:213-238.

Croft, W. 2001. *Radical construction grammar: Syntactic theory in typological perspective.* Oxford: Oxford University Press.

Croft, W., and A. Cruise. 2004. *Cognitive linguistics.* Cambridge: Cambridge University Press.

de Bot, K., W. Lowie, and M. Verspoor. 2007. A dynamic systems theory to second language acquisition. *Bilingualism: Language and Cognition* 10:7-21.

Dietrich, R., W. Klein, and C. Noyau, eds. 1995. *The acquisition of temporality in a second language.* Amsterdam: John Benjamins.

Ebbinghaus, H. 1885. *Memory: A contribution to experimental psychology.* Translated by H. A. R. C. E. B. (1913). New York: Teachers College, Columbia.

Elio, Renee, and J. R. Anderson. 1981. The effects of category generalizations and instance similarity on schema abstraction. *Journal of Experimental Psychology: Human Learning & Memory* 7 (6):397-417.

———. 1984. The effects of information order and learning mode on schema abstraction. *Memory & Cognition* 12 (1):20-30.

Ellis, N. C. 1998. Emergentism, connectionism and language learning. *Language Learning* 48 (4):631-664.

———. 2002. Frequency effects in language processing: A review with implications for theories of implicit and explicit language acquisition. *Studies in Second Language Acquisition* 24 (2):143-188.

———. 2002. Reflections on frequency effects in language processing. *Studies in Second Language Acquisition* 24 (2):297-339.

———. 2003. Constructions, chunking, and connectionism: The emergence of second language structure. In *Handbook of second language acquisition*, edited by C. Doughty and M. H. Long. Oxford: Blackwell.

———. 2006. Language acquisition as rational contingency learning. *Applied Linguistics* 27 (1):1-24.

———. 2006. Selective attention and transfer phenomena in SLA: Contingency, cue competition, salience, interference, overshadowing, blocking, and perceptual learning. *Applied Linguistics* 27 (2):1-31.

———. 2007. Blocking and Learned Attention in Language Acquisition. *CogSci 2007, Proceedings of the Twenty Ninth Cognitive Science Conference. Nashville, Tennesse, August 1-4, 2007.*

———. 2008. The dynamics of language use, language change, and first and second language acquisition. *Modern Language Journal* 41 (3):232-249.

———. 2008. The Psycholinguistics of the Interaction Hypothesis. In *Multiple perspectives on interaction in SLA: Second language research in honor of Susan M. Gass*, edited by A. Mackey and C. Polio. New York: Routledge.

———. 2008. Usage-based and form-focused language acquisition: The associative learning of constructions, learned-attention, and the limited L2 endstate. In *Handbook of cognitive linguistics and second language acquisition*, edited by P. Robinson and N. C. Ellis. London: Routledge.

Ellis, N. C., and T. Cadierno. 2009. Constructing a second language. *Annual Review of Cognitive Linguistics* 7 (Special section).

Ellis, N. C., and F. Ferreira-Junior. 2009. Construction Learning as a function of Frequency, Frequency Distribution, and Function *Modern Language Journal* 93:370-386.

———. 2009. Constructions and their acquisition: Islands and the distinctiveness of their occupancy. *Annual Review of Cognitive Linguistics*:111-139.

Ellis, N. C., and D. Larsen-Freeman. 2006. Language emergence: Implications for Applied Linguistics. *Applied Linguistics* 27 (4):whole issue.

———. 2006. Language emergence: Implications for Applied Linguistics (Introduction to the Special Issue). *Applied Linguistics* 27 (4):558-589.

———. 2009. Constructing a second language: Analyses and computational simulations of the emergence of linguistic constructions from usage. . *Language Learning* 59 (Supplement 1):93-128.

———. 2009. Language as a Complex Adaptive System (Special Issue). *Language Learning* 59: Supplement 1.

Ellis, N. C., and N. Sagarra. 2010. Learned Attention Effects in L2 Temporal Reference: The First Hour and the Next Eight Semesters. *Language Learning* 60 Supplement 1.

———. 2010. The bounds of adult language acquisition: Blocking and learned attention *Studies in Second Language Acquisition* 32 (4).

Ellis, N. C., and R. Schmidt. 1998. Rules or associations in the acquisition of morphology? The frequency by regularity interaction in human and PDP learning of morphosyntax. *Language & Cognitive Processes* 13 (2&3):307-336.

Ellis, N. C., R. Simpson-Vlach, and C. Maynard. 2008. Formulaic Language in Native and Second-Language Speakers: Psycholinguistics, Corpus Linguistics, and TESOL. *TESOL Quarterly* 42 (3):375-396.

Elman, J. L., E. A. Bates, M. H. Johnson, A. Karmiloff-Smith, D. Parisi, and K. Plunkett. 1996. *Rethinking innateness: A connectionist perspective on development.* Cambridge, MA: MIT Press.

Evert, S. 2005. The Statistics of Word Cooccurrences: Word Pairs and Collocations, University of Stuttgart, Stuttgart.

Ferrer i Cancho, R., and R. V. Solé. 2001. The small world of human language. *Proceedings of the Royal Society of London, B.* 268:2261-2265.

———. 2003. Least effort and the origins of scaling in human language. *PNAS* 100:788-791.

Ferrer i Cancho, R., R. V. Solé, and R. Köhler. 2004. Patterns in syntactic dependency networks. *Physical Review* E69:0519151-0519158.

Goldberg, A. E. 1995. *Constructions: A construction grammar approach to argument structure.* Chicago: University of Chicago Press.

———. 2003. Constructions: a new theoretical approach to language. *Trends in Cognitive Science* 7:219-224.

———. 2006. *Constructions at work: The nature of generalization in language.* Oxford: Oxford University Press.

Goldberg, A. E., D. M. Casenhiser, and N. Sethuraman. 2004. Learning argument structure generalizations. *Cognitive Linguistics* 15:289–316.

Goldschneider, J. M., and R. DeKeyser. 2001. Explaining the "natural order of L2 morpheme acquisition" in English: A meta-analysis of multiple determinants. *Language Learning* 51:1-50.

Gries, S. Th. 2008. Corpus-based methods in analyses of SLA data In *Handbook of cognitive linguistics and second language acquisition*, edited by P. Robinson and N. C. Ellis. London: Routledge.

Gries, S. Th., and A. Stefanowitsch. 2004. Extending collostructional analysis: a corpus-based perspective on 'alternations'. *International Journal of Corpus Linguistics* 9:97-129.

Gries, S. Th., and S. Wulff. 2005. Do foreign language learners also have constructions? Evidence from priming, sorting, and corpora. *Annual Review of Cognitive Linguistics* 3:182-200.

Harnad, S., ed. 1987. *Categorical perception: The groundwork of cognition*. New York: Cambridge University Press.

Hoey, M. P. 2005. *Lexical priming: A new theory of words and language*. London: Routledge.

Holland, J. H. 1995. *Hidden order: How adaption builds complexity*. Reading: Addison-Wesley.

———. 1998. *Emergence: From chaos to order*. Oxford: Oxford University Press.

James, W. 1890. *The principles of psychology*. Vol. 1. New York: Holt.

Jurafsky, D. 2002. Probabilistic Modeling in Psycholinguistics: Linguistic Comprehension and Production. In *Probabilistic Linguistics*, edited by R. Bod, J. Hay and S. Jannedy. Harvard, MA: MIT Press.

Jurafsky, D., and J. H. Martin. 2000. *Speech and language processing: An introduction to natural language processing, speech recognition, and computational linguistics*. Englewood Cliffs, NJ: Prentice-Hall.

Kello, C. T., and B. C. Beltz. 2009. Scale-free networks in phonological and orthographic wordform lexicons. In *Approaches to Phonological Complexity*, edited by I. Chitoran, C. Coupé, E. Marsico and F. Pellegrino. Berlin: Mouton de Gruyter.

Kello, C. T., G. D. A. Brown, R. Ferrer-i-Cancho, J. G. Holden, K. Linkenkaer-Hansen, T. Rhodes, and G. C. Van Orden. 2010. Scaling laws in cognitive sciences. *Trends in Cognitive Science* 14:223-232.

Lakoff, G. 1987. *Women, fire, and dangerous things: What categories reveal about the mind*. Chicago: University of Chicago Press.

Langacker, R. W. 1987. *Foundations of cognitive grammar: Vol. 1. Theoretical prerequisites*. Stanford, CA: Stanford University Press.

Larsen-Freeman, D. 1997. Chaos/complexity science and second language acquisition. *Applied Linguistics* 18:141-165.

Larsen-Freeman, D., and L. Cameron. 2008. *Complex systems and applied linguistics*. Oxford: Oxford University Press.

Lieberman, E., J-B. Michel, J. Jackson, T. Tang, and M. A. Nowak. 2007. Quantifying the evolutionary dynamics of language. *Nature* 449 (7163):713--716.

MacWhinney, B. 1987. Applying the Competition Model to bilingualism. *Applied Psycholinguistics* 8 (4):315-327.

———. 1987. The Competition Model. In *Mechanisms of language acquisition*, edited by B. MacWhinney. Hillsdale, NJ: Erlbaum.

———. 1997. Second language acquisition and the Competition Model. In *Tutorials in bilingualism: Psycholinguistic perspectives*, edited by A. M. B. De Groot and J. F. Kroll. Mahwah, NJ: Lawrence Erlbaum Associates.

———. 2001. The competition model: The input, the context, and the brain. In *Cognition and second language instruction*, edited by P. Robinson. New York: Cambridge University Press.

———, ed. 1999. *The emergence of language*. Hillsdale, NJ: Erlbaum.

Manin, D. Y. 2008. Zipf's Law and Avoidance of Excessive Synonymy *Cognitive Science* 32:1075-1098.

McDonough, K. 2006. Interaction and syntactic priming: English L2 speakers' production of dative constructions. *Studies in Second Language Acquisition* 28:179-207.

McDonough, K., and A. Mackey. 2006. Responses to recasts: Repetitions, primed production and linguistic development. *Language Learning* 56:693-720.

McDonough, K., and P. Trofimovich. 2008. *Using priming methods in second language research*. London: Routledge.

Nation, P. 2001. *Learning vocabulary in another language*. Cambridge: Cambridge University Press.

Newell, A. 1990. *Unified theories of cognition*. Cambridge, MA: Harvard University Press.

Newman, M. 2003. The structure and function of complex networks. *SIAM Review* 45:167-256.

———. 2005. Power laws, Pareto distributions and Zipf's law. *Contemporary Physics* 46:323–351

Ninio, A. 1999. Pathbreaking verbs in syntactic development and the question of prototypical transitivity. *Journal of Child Language* 26:619-653.

———. 2006. *Language and the learning curve: A new theory of syntactic development*. Oxford: Oxford University Press.

O'Donnell, M., and N.C. Ellis. 2009. Measuring formulaic language in corpora from the perspective of Language as a Complex System. In *5th Corpus Linguistics Conference*. University of Liverpool.

O'Donnell, M., and N.C. Ellis. 2010. Towards an Inventory of English Verb Argument Constructions. *Proceedings of the 11th Annual Conference of the North American Chapter of the Association for Computational Linguistics, Los Angeles*.

Pagel, M., Q. D. Atkinson, and A. Meade. 2007. Frequency of word-use predicts rates of lexical evolution throughout Indo-European history. *Nature* 449 (7163):717--720.

Pickering, M. J. 2006. The dance of dialogue. *The Psychologist* 19:734-737.

Pickering, M. J., and V. S. Ferreira. 2008. Structural priming: A critical review. *Psychological Bulletin* 134:427-459.

Pickering, M. J., and S. C. Garrod. 2006. Alignment as the basis for successful communication. *Research on Language and Computation* 4:203-228.

Pinker, S. 1989. *Learnability and cognition: The acquisition of argument structure*. Cambridge, MA: Bradford Books.

Port, R. F., and T. Van Gelder. 1995. *Mind as motion: Explorations in the dynamics of cognition*. Boston MA: MIT Press.

Posner, M. I., and S. W. Keele. 1970. Retention of abstract ideas. *Journal of Experimental Psychology* 83:304-308.

Rescorla, R. A. 1968. Probability of shock in the presence and absence of CS in fear conditioning. *Journal of Comparative and Physiological Psychology* 66:1-5.

Rescorla, R. A., and A. R. Wagner. 1972. A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In *Classical conditioning II: Current theory and research*, edited by A. H. Black and W. F. Prokasy. New York: Appleton-Century-Crofts.

Robinson, P., and N.C. Ellis, eds. 2008. *A handbook of cognitive linguistics and second language acquisition*. London: Routledge.

Rosch, E., and C. B. Mervis. 1975. Cognitive representations of semantic categories. *Journal of Experimental Psychology: General* 104:192-233.

Rosch, E., C. B. Mervis, W. D. Gray, D. M. Johnson, and P. Boyes-Braem. 1976. Basic objects in natural categories. *Cognitive Psychology* 8:382-439.

Saussure, F. de. 1916. *Cours de linguistique générale*. Translated by Roy Harris. London: Duckworth.

Shanks, D. R. 1995. *The psychology of associative learning*. New York: Cambridge University Press.

Slobin, D. I. 1997. The origins of grammaticizable notions: Beyond the individual mind. In *The crosslinguistic study of language acquisition*, edited by D. I. Slobin. Mahwah, NJ: Erlbaum.

Solé, R. V., B. Murtra, S. Valverde, and L. Steels. 2005. Language Networks: their structure, function and evolution. *Trends in Cognitive Sciences* 12.

Spencer, J. P., M. S. C. Thomas, and J. L. McClelland, eds. 2009. *Toward a unified theory of development: Connectionism and dynamic systems theory re-considered*. New York: Oxford.

Spivey, M. 2006. *The continuity of mind*. Oxford: Oxford University Press.

Stefanowitsch, A., and S. Th. Gries. 2003. Collostructions: Investigating the interaction between words and constructions. *International Journal of Corpus Linguistics* 8:209-43.

Swales, J. M. 1990. *Genre analysis : English in academic and research settings*. Cambridge: Cambridge University Press.

Taylor, J. R. 1998. Syntactic constructions as prototype categories. In *The new psychology of language: Cognitive and functional approaches to language structure*, edited by M. Tomasello. Mahwah, NJ: Erlbaum.

Tennenbaum, J. 2005. The large-scale structure of semantic networks: statistical analyses and a model of semantic growth. *Cognitive Science* 29:41-78.

Tomasello, M. 2003. *Constructing a language*. Boston, MA: Harvard University Press.

———, ed. 1998. *The new psychology of language: Cognitive and functional approaches to language structure*. Mahwah, NJ: Erlbaum.

Tversky, A. 1977. Features of similarity. *Psychological Review* 84:327-352.

van Geert, P. 1991. A dynamic systems model of cognitive and language growth. *Psychological Review* 98 (3-53).

Van Patten, B. 2006. Input processing. In *Theories in second language acquisition: An introduction*, edited by B. Van Patten and J. Williams. Mahwah, NJ: Lawrence Erlbaum.

Wittgenstein, L. 1953. *Philosophical investigations*. Translated by G. E. M. Anscombe. Oxford: Blackwell.

Wulff, S., N. C. Ellis, U. Römer, K. Bardovi-Harlig, and C. LeBlanc. 2009. The acquisition of tense-aspect: Converging evidence from corpora, cognition, and learner constructions. *Modern Language Journal* 93:354-369.

Zipf, G. K. 1935. *The psycho-biology of language: An introduction to dynamic philology*. Cambridge, MA: The M.I.T. Press.

———. 1949. *Human behaviour and the principle of least effort: An introduction to human ecology*. Cambridge, MA: Addison-Wesley.